Defense Technical Information Center
Compilation Part Notice

ADP012336

TITLE: Agents On the Semantic Object Web: Information Management for
Coalition Operations

DISTRIBUTION: Approved for public release, distribution unlimited
Availability: Hard copy only.

This paper is part of the following report:

TITLE: KSCO 2002: Second International Conference on Knowledge
Systems for Coalition Operations

To order the complete compilation report, use: ADA402533

The component part is provided here to allow users access to individually authored sections
of proceedings, annals, symposia, etc. However, the component should be considered within
the context of the overall compilation report and not as a stand-alone technical report.

The following component part numbers comprise the compilation report:
ADP012330 thru ADP012354

Agents On The Semantic Object Web:
Information Management for Coalition Operations

G. Edwards, B. Kettler, K. Olin, and B. Tsurutani

ISX Corporation
4301 N Fairfax Drive, Suite 370
Arlington, Virginia 22203
gedwards@isx.com

Abstract. Coalition operations pose clear challenges for information sharing and for the integration of disparate information processes. ISX Corporation has been investigating the combined use of agent-based approaches to implement Information Management Agents™ operating on a semantically organized Semantic Object Web™ to provide highly flexible means of delivering information services within a large, distributed and diverse enterprise. The techniques described demonstrate their particular suitability to rapidly standing up a "come as you are" organization of coalition partners whose information protocols, requirements, and processes might vary widely. The use of a Semantic Object Web provides an agent-navigable information substrate which can be used by service-provider agents to discovery relevant information or services, to map their own content into a shareable semantic space, and to exploit available content and services. Using agent-based techniques, the approaches described provide the ability to deconstruct information requirements to guide matching of source / consumer relationships within the enterprise, and then to compose, aggregate, and transform information from various sources to meet those requirements. In this vein, the techniques provide services registration, matching and brokerage, agent facilitation and control, and semantic matching and transformation necessary to compose the right information for the coalition force member. In addition, this work also addresses the dissemination and delivery of information. Information Management Agents provide the means to constrain and guide information access and delivery within the coalition means to implement selective information management business rules and policy server-imposed access and publication restriction. Together, these technologies promise the capability to support interoperability across coalition organizations while maintaining necessary policy and process-based constraints on information access and dissemination constraints.

1   Introduction

Information technology to support integrated coalition operations in future military environments poses a two-edged opportunity. Effective coalition operations will rely on information technology as a key enabler. Coalition operations are all about bringing together a diverse set of organizations, each with their own capabilities, processes, and infrastructure, and forming them into a working enterprise tailored to the military operation at hand. Information technology-enabled sharing and coordination are at the heart of successful coalition operations. At the same time, coalitions are not seamless organizations. Effective sharing and coordination, to occur at all, must also mean acceptable sharing and coordination. In the real world, information technology must also enforce constraints on what information can be shared and what services can be provided, under what conditions, and to whom.

In this paper, we will explore selected key facets of the problem of integrating diverse command and control organizations into a unified enterprise in the face of such constraints. Our discussion will focus on the interplay of three key problem elements: interoperability, or how information can be practically exchanged and shared across coalition elements; process, focusing on how information technology can enforce workable C2 enterprise "business relationships" between organizations in the coalition; and policy, or how information technology can enforce organizationally-imposed rules that let coalition partners share and coordinate within constraints they impose on their own information assets. Our technology discussion will focus on primarily on the roles of two key technologies. The first of these technologies, the Semantic Object Web ™, provides an ontologically grounded means of indexing diverse coalition information sources into a single integrated information space. The second technology approach exploits self-organizing collections of software agents, capable of querying and navigating the Semantic Object Web, as tools to deliver both information access and delivery services, and to implement information management constraints, processes and policies. We believe these technologies, in concert, provide an interesting approach to dealing with key problems in the coalition enterprise space. While we are not proposing a comprehensive architecture or solution to the coalition management problem, we believe the ideas expressed here, and the initial experimental work we have undertaken, provide some interesting and potentially powerful approaches.

2   Operational Demands

Information is the commodity that enables coordinated and effective military operations. Effective exploitation of that information is key to successful operations. One of the challenges for a military coalition commander and his staff is to organize a command and control (C2) enterprise built around the requirements and characteristics of a specific military operation. This enterprise must stand up quickly to put working operational capabilities into place. And increasingly, in a world where a broad community of international partners share a global interest in regional stability, this enterprise will include coalition partners. Creating a working command and control enterprise in this context poses a wide range of issues, from simple interoperability issues to complex mechanisms for services discovery and negotiation of roles. Unlike the past century, these coalitions will not be limited to NATO-like organizations with long-standing alliances and deeply rooted operational practice refined and coordinated in numerous exercises and operations. The problem for information technology is to take a complex set of extant command and control capabilities, often with partners whose command and control organizations have never worked with each other before (and were not designed to do so), and to integrate and shape those capabilities into an operationally-tailored working enterprise that enables both interoperability and coordination.

Information technology must facilitate the commander's ability to exercise control over the information flows within this enterprise, and must also meet all of the information management constraints imposed by participating organizations. Modern coalitions will be composed of more loosely associated groups of nations, each with its own level of commitment to the coalition, each with its own agenda (which will certainly only partially overlap interests of its coalition partners), and each engaged in a limited role within the operational enterprise. Certainly, many concerns about information management, dissemination, access, etc. are driven by security concerns. Loose coalition partnerships may include nations whose partnership is very limited, such as Pakistan and India both participating in a counter-terrorism operation. The level of participation offered by such partners in intelligence and C2 processes will depend in part on their ability to protect and manage their own assets. This may often mean protecting intelligence sources and methods, information gathering capabilities, and battlefield capabilities as well as specific situation assessment or plan products, while making some of those information elements available to coalition command to drive the command and control process. A related concern is the level of security maintained once content leaves, for example, US command and control enclaves and is posted within enclaves which might have less strident security features in place. Information assurance beyond access control must be considered an equally important source of constraints on how information is handled. From an information assurance point of view, where information is coming from, and how it was processed to produce decision-level information, is as important as what is done with the information or who has access to the product. Finally, because of the complexity of a large coalition command and control enterprise, the business processes of the enterprise itself must be enforced. These business processes define what each coalition organization is responsible (and allowed) to provide to the overall enterprise as well as what they are allowed to demand from the enterprise. These additional constraints on information access, dissemination, and processing are necessary to maintain a coherent and stable set of processes that can effectively perform command and control information processing and decision making activities without grinding to a halt or making conflicting or poorly informed decisions.

In short, the issues we have been addressing in our research are driven by the need to stand up a working command and control environment with supporting technology that integrates diverse information producer and consumer organizations. We are much less concerned about how an individual operator or staff member accesses information, and much more concerned with how organizations that do not typically work in concert pull together into a working enterprise. We believe the best conceptual example of this problem is the Joint Battlespace Infosphere model, developed in 1997 by a U.S. Air Force Scientific Advisory Board ad hoc study committee. This concept is the basis for several related US Air Force sponsored research initiatives, and provides the conceptual framework in which many of our working examples and problems are framed (some under US Air Force / AFRL sponsorship, some under independent sponsorship). We have also been fortunate to explore related problems in information discovery and information sharing across enclave boundaries in the military intelligence community. As we look across these various operational models and the use cases that derive from them, we find ourselves consistently facing three key problem areas: interoperability, enforcement of information management policy, and enforcement of enterprise information processes. In looking at mechanisms to address these issues, we have found a high degree of utility from two particular classes of technology: Semantic Object Webs, and Information Management Agents.

3   The Semantic Object Web ™ – Organizing Content with Semantic Underpinnings

One very fruitful area of technology exploration at ISX has focused on the exploitation of semantic underpinnings to aid in information discovery and interoperability. Across many domains, including command and control, highly structured databases and standard publication formats (not to mention paper documents) are giving way to rapid publication and update of less structured content augmented with XML-based metadata and embedded markup. XML and derivative tag-language based interoperability is becoming a de facto standard mechanism for externalizing information, for packaging information for transport to targeted applications, and for allowing content to be exploited in ways often unanticipated by the information producer. These approaches are driven by the recognition that, unlike interoperability based solely on shared data models and common format shared repositories. By adding a layer of semantic mapping between information (content and services) and a shared domain ontology, these approaches have demonstrated advantages of more robust interoperability between diverse information providers and consumers. Just as HTML provided the means to express how information format should be interpreted by a browser, XML-based domain languages have provided the means to guide how an information exploitation capability should understand the content provided by an information source. Increasingly, these techniques are being exploited not only to provide enhanced interoperability, but to provide a model of the information content or services provided by a source, or the information requirements of a consumer. More advanced semantic markup languages like the DARPA Agent Markup Language (DAML), combined with an emerging generation of DAML-based tools, provide the mechanisms to develop and exploit embedded markup and metadata for both unstructured sources (such as analysis reports) and highly structured sources (such as databases). These tools give us the ability to not only understand the semantic mappings necessary to enable interoperability, but provide the means for humans or software agents to browse large collections of information elements and to explore complex links and relationships. On the information integration side, this ability to build a highly connect space of semantic links provides the means to integrate information that is divers in both content and structure. On the exploitation side, this enables much more powerful information discovery and retrieval as well as the advantages of interoperability noted above.

Rapidly escalating trends toward the exploitation of semantic markup and metadata offer potentially powerful approaches that promise to meet some of the most challenging problems in the coalition environment. First, consider the nature of future coalition operational problems. Traditional fights with traditional enemies are becoming increasingly infrequent. New kinds of battles, against non-traditional adversaries, highly asymmetric threats, and non-traditional battlefields are becoming increasingly frequent. Many of the information processes and products of today's command and control environments will be hard pressed to meet the demands of these new classes of problems. For example, when the US and its coalition partners take on Al Qaeda operations worldwide, new information processes will be needed. The products of intelligence gathering and operational planning processes will not be able to take days or weeks to massage data into standardized formats, resolve uncertainties and contradictions, and publish "authoritative" databases. Nor will operational plans be able to rely on standardized publications at regular intervals to keep everyone "on the plan." Reaction cycles are getting too short, and information is getting stale too quickly. Information will need to come from whatever sources are available, and will have to be updated and shared in increasingly raw form, with tools to help information analysts identify and augment key content, and for consumers to quickly find and extract the content they need. Second, consider the changing nature of a real-world coalition environment and the roles of coalition players. In the past, coalition has often meant a U.S. run operation, where other players offer cooperation and support, even battlefield resources, but the primary challenge was coordination of forces. Increasingly, we will face problems where timely reaction to intelligence dominates over our ability to coordinated force as the key to the winning formula, and where organizations across the coalition play important roles in keeping the flow of global intelligence connected to the coordination of forces. In these coalition problems, the flexibility of semantically grounded representation offers

Figure 1: Semantic Object Web example

44

the potential key to information discovery and sharing between very diverse organization, each with their own tools, their own representations, and their own business processes. Such representations promise to form the basis for interoperability and exploitation of information products without detailed, pre-defined agreements on definitions of specific information product formats. This technology is moving the military command and control world rapidly from a data format-based model of interoperability to a semantics-based model of interoperability.

In our approach to the exploitation of semantic tag languages (in our case DAML being the ontologically grounded language of choice), we are building on the notion of the Semantic Object Web. Recognizing the Web as today's best example of a large, shared information structure, Tim Berners-Lee coined the term "Semantic Web" to describe the evolutionary model of a future Web that allowed machines, not just people, to exploit content and services. Using an expressive semantic markup that makes content, metadata, and services available and understandable, and using tools that exploit shared semantics represented as ontologies, user applications or automated software agents could both publish and consume information from the Web. DAML represents the current results of a body of DARPA-funded researchers pursuing the notion of the Semantic Web by providing the semantic standards to extend simple XML. We have found the notion of the Semantic Web a powerful metaphor for the kinds of information exploitation needed to support future command and control environments. The Semantic Object Web implements some of our ideas about how to take advantage of a practical Semantic Web implementation for distributed enterprise information sharing.

The Semantic Object Web takes an object-oriented approach to modeling the available content in a Semantic Web. This model consists of a semantic network of objects, each of which represents an entity whose type is defined in a domain ontology, with links to other entity objects based on relationships also defined in a domain ontology. The Semantic Object Web is built by processing the available markup-based metadata and content markup, along with XML-externalized database schema and content (schema to provide exploitable structure, and content to help resolve enitiy and relationship unification). As each element of information is processed in the construction of the Semantic Object Web, an inference engine (currently the PARKA system, developed by University of Maryland) maps the new content into the existing model. The inference engine operates over the new content, existing model, and a pre-defined set of ontological specifications and mappings to unify references to entities in the domain, to recognize information that implies relationships between these entities, and to identify references to content about the attributes of the entities. Each object maintains a set of pointers to source information related to its definition and its attributes, and each relationship object is used to maintain a set of pointers to the source material supporting each inferred relationship. These objects and link references can be stored in one or more distributable databases for efficient exploitation by humans with appropriate query and navigation tools, or by software agents capable of doing an initial query and then "walking the structures" to search for complex object/relationship substructures of interest. The Semantic Object Web can be thought of as an efficient agent-exploitable index into the content of a Semantic Web. Having found the right entities and relationships in the "index" model, an agent can pursue the specific links to supporting source material to extract and deliver necessary content.

Our initial experiments offer promising results to address the problems of information integration, discovery, retrieval, and interoperability across organizational boundaries. We have demonstrated the ability to integrate information metadata and markup from unique sub-domain ontologies into a single Semantic Object Web by creating fairly simple mappings between ontologies. And while our initial experiments were based on human query and navigation tools, we believe that agent-based exploitation of these structures are quite practical, as we hope to demonstrate in future work. Given such a capability, we can address some key problems in coalition interoperability by breaking down semantic barriers, providing mechanisms for the discovery of relevant information, and providing and extensible information architecture for a large, diverse, and distributed coalition community.

4   Agent-Based Information Management

Another focus area for coalition-relevant research at ISX is in the use of software Information Management Agents (supported by the COABS Grid services) as rapidly and dynamically composable components for both accessing information and for implementing information management strategies. Considering some of the key aspects of information management problems facing the coalition enterprise, we believe such approaches offer a useful and necessary capability. We have observed that these approaches can provide services to help information requestors find and access relevant information, transform it into useful abstractions or formats, deliver it on demand, and monitor for relevant changes.

Our exploration of agent-based approaches to information management stems from our recognition of some key problems encountered when trying to interoperate across organizational information processes, each with its own semantics for information representation. The most apparent problem is the simple semantic data model mismatch between organizational models. When an organization establishes a flow of information into its processes, it must

be able to exploit that information with tools that were designed to work with a specific data representation, which captures a specific level of abstraction and aggregation of source information. To establish a flow of useable information into an organization's process, several steps may be necessary to bridge this semantic mismatch. First, the consumer must be able to find the kind of information sources it wants. This might not be a simple mapping, as the available data might be represented using different semantics, and may even need to be composed out of various pieces of information aggregated or assembled into the needed product. Information sources must be identified. Abstractions and representation in the request must be deconstructed to find a workable mapping between requestor and provider information elements. Queries must be decomposed to match those mappings, and the results passed through the right aggregation and fusion processes to create the desired content. Finally, the resulting information must be transformed into the right abstraction and format to enable interoperability with the consumer's tools.

A second class of problem involves the enforcement of constraints on these information flows. Here, our concern is on enforcing various restrictions on what kinds of information or services an organization can exploit, and on what kinds of information content and services that organization can make available. Today, such constraints are largely enforced by limiting access to entire broad classes of information, such as restricting access to systems which operate at certain levels of classification, or limiting access to information from certain sources. A more desirable capability would be to apply policies that define more specific information protection and process management objectives of the organizations involved, and to provide an automated mechanism to make sure that these policies are properly applied. Such policies could express the intent to protect certain sources or methods, to limit dissemination of certain sensitive intelligence or operational plans, and to enforce specific role-based information service constraints on partner organizations. Ideally, consideration of such policies would be considered as part of the process of establishing information flows between organizations, and their application would allow some flexibility in meeting policy requirements while supporting key information requirements. For example, a policy to protect an information source might be implementable (perhaps with human oversight and approval) by delivering needed content only in an aggregate form, where source-specific relationships are abstracted away. Our research is focused on the role Information Management Agents can play in both implementing interoperability bridges between disparate sources and consumers, and in implementing mechanisms to enforce information management policies and processes.

Work at ISX has explored several specific classes of agent-based information management problems relevant to the objectives described above. In the first, our goal was to de-couple information consumers from information sources, and to demonstrate that software agents could enable more loosely coupled modes of interoperability. In our experiments, we provided a mechanism to handle information requests from consumers which would typically access some shared, common-format repository. Instead, we introduced a collection of software agents designed to implement these information requests by dynamically organizing various agent-base functions operating across a range of information sources. Facilitation agents implemented selected classes of information requests by soliciting and organizing the activities of various other functional agents. Some of these agents provided the ability to decompose the information request into finer grained requests to match available information source services, while others provided the ability to re-assemble or re-aggregate the component results to satisfy the information request. Other agents provided translation services, mapping information requests into the right semantics to match information sources, and translating the resulting information into the form needed by the requestor. And of course, a collection of service brokering and matching agents were required to help the facilitators find the right agents to compose into a working access service.
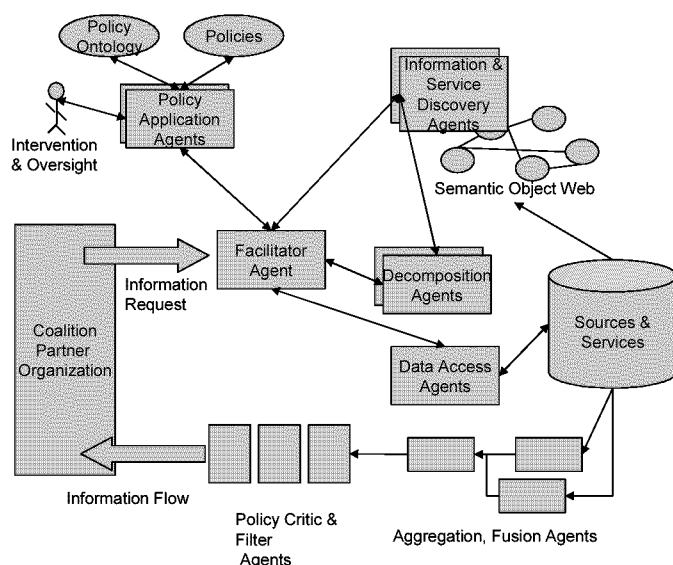
In a second research project, funded by the USAF / AFRL Joint Battlespace Infosphere (JBI) project, we use a similar agent-based approach to augment an existing publish and subscribe service on the JBI platform. In this work, we provided a lightweight agent framework to construct "fuselet" agents capable of simple information aggregation and transformation operations. Given a consumer request to the publish-subscribe mechanism which fails to match any available information publication source, the failed request is handed over to the "fuselet" mechanism. This mechanism is capable of decomposing the request, matching the decomposed elements to available

Figure 2: Information Management Agents concept

46

subscriptions, establishing those subscriptions, and re-composing the results for delivery to the requestor. The re-composition might be as simple performing simple set aggregation or counting, or might involve a lightweight fusion operator to combine multiple inputs into the desired information. In the end, transformational agents turn the acquired information into the desired abstraction and format for delivery as a new "subscription source."

In both of these initial experiments, we were able to demonstrate that simple agent-based information management functions could be automatically composed on demand to implement fairly complex information management tasks. While these experiments focused on the interoperability-oriented aspects of decomposing requests, gathering data, recomposing results, and transforming semantics, they illustrate a useful level of complexity and robustness we believe could be applicable in coalition- related information policy and process management problems. Currently proposed extensions to this work will attempt to generalize these agents to go beyond information access, and to directly implement constraints imposed on information access and dissemination by organizational policy. In these experiments, we intend to provide facilitation agents that serve as critics on information subscription or publication requests from a producer or consumer organization to the JBI platform. By checking these requests against a hierarchy of organizational policy models and a process model for the command and control enterprise, these agents can either reject un-allowed requests, or can assemble a collection of information transformation agents that can filter or abstract information to meet policy-imposed or process-imposed constraints.

5   The Future: Agents On The Semantic Object Web

In our future research, we hope to begin to put together many of the pieces described in the previous sections to provide a more comprehensive experimental model for coalition information management. Our focus will be the exploitation of the Semantic Object Web by software agents designed to implement both interoperability and information access and dissemination management services.

To support interoperability, we intend to extend the notions of agent-based decomposition, search, and retrieval to exploit semantic structure. Information requests today depend wholly on shared data models, and queries typically request known data structures or content elements. Today, we can provide the ability to query the Semantic Object Web using more general ontology-based queries, such as "Give me all the entities of type Threat" and expect to find matches to various sub-types and specializations of threat. In the near future, we anticipate asking agents to handle much more complex queries, like "Find any link between John Hatfield and David McCoy," or more relevant to our domain, "Find any hostile organizations or forces that might resist SOF team deployment in Village X." And we will expect these agents to deliver the content, pulled from the best available sources, delivered in a form we can use. Such capabilities offer the promise of much richer interaction between organizational elements, based on better indexing and better retrieval of all available information despite organizational semantic barriers.

To support better information management, we intend to explore the use of Information Management Agent-based implementation of information access and dissemination control policies and enterprise process policies represented as Semantic Object Web structures. While we are very interested in ongoing work in policy servers and policy ontology, we intend to specifically explore the power of semantic indexing and retrieval mechanisms to allow agents to quickly find relevant policy elements, and to intelligently apply those policies to specific information instances. Using these techniques, we believe that agents will be able to consider metadata about specific object abstractions and relationships, individual attributes, and specific sources and source types as the basis of dissemination constraints. Given this capability, a policy might be expressed in fairly abstract terms that address categories or types of sources, content attributes and relationships, representational abstractions. Agent critics could use the ontology-base inferences in the Semantic Object Web to compare these policies to actual instances of information access or publication requests.

In summary, our initial research activities have shown both the potential power and practicality of both Semantic Object Web and Information Integration Agent technologies to address key problems of information interoperability and management in domains like coalition operations. This early work features successful experiments with problems whose characteristics and complexities match those we anticipate for more comprehensive coalition problems. However, we also recognize that we have only scratched the surface of these problems. We expect to make significant steps forward by transitioning our agent capabilities to exploit our semantic information substrate both for content and services exploitation, and for exploitation of information management policy. We also are eager to identify opportunities for joint experiments with related technology research to help expend these ideas and their application to real problems.

6 References

Berners-Lee, Tim. 1999. *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by its Inventor*. HarperCollins.

Berners-Lee, Tim, James Hendler, and Ora Lassila. 2001. The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American* (May, 2001).

DARPA Agent Markup Language (**DAML**) Project and Language: http://www.daml.org

Parka Knowledge Representation System: http://www.cs.umd.edu/projects/plus/Parka/

Simple Ontology HTML Extensions (**SHOE**) Project and Language: http://www.cs.umd.edu/projects/plus/SHOE/

Resource Description Framework (**RDF**): http://www.w3.org/RDF/ (RDF Schemas: http://www.w3.org/TR/2000/CR-rdf-schema-20000327/)

World Wide Web Consortium's Semantic Web Activity: http://www.w3c.org/2001/sw/

World Wide Web Consortium's Web-Ontology Working Group: http://www.w3c.org/2001/sw/WebOnt/

Heflin, Jeff. 2001. Towards the Semantic Web: Knowledge Representation in a Dynamic Distributed Environment. Ph.D. diss., Dept. of Computer Science, Univ. of Maryland at College Park.

Fuselet Project Final Report, ISX Corp., Jan 2002

Marmelstein, R., *Force Template White Paper*, (Notes from the AFRL Minnowbrook Joint Battlespace Infosphere workshop, Force Template Working Group), AFRL Nov 2001

Linderman, M*., The JBI Platform*, Briefing to the Minnowbrook Joint Battlespace Infosphere workshop, Nov. 2001

McCarthy, J. (ed.), *Building the Joint Battlespace Infospher*e, US Air Force Scientific Advisory Board, Summer Study Report, Dec. 1999

McCarthy, J. (ed.), *Information Management to Support the Warrior*, US Air Force Scientific Advisory Board, Summer Study Report, Nov. 1998